November 30, 2011

VIA E-MAIL

Dr. Ruth Lunn
Director, Office of the Report on Carcinogens
Division of the National Toxicology Program
National Institute of Environmental Health Sciences
P.O. Box 12203, MD-K2-14
Research Triangle Park, North Carolina 27709
lunn@niehs.nih.gov

Re: Proposed Revisions to the Review Process for the Report on Carcinogens

Dear Dr. Lunn:

The American Chemistry Council ("ACC") appreciates the opportunity to comment on the National Toxicology Program's ("NTP") proposed revisions to the review process ("Review Process") for its Report on Carcinogens ("RoC"). 76 Fed. Reg. 67200 (Oct. 31, 2011). ACC has a longstanding interest in the RoC and its review process, and commented extensively on NTP's December 2003 proposed RoC review process.

ACC commends NTP for evaluating needed revisions to the Review Process in order to enhance the transparency and efficacy of NTP's RoC. However, ACC is disappointed that NTP's stated objective is merely to "maintain," rather than improve, those elements that relate to the scientific integrity of the Review Process. As set forth in the enclosed comments, ACC believes that in addition to improving the peer review and public comment aspects of the Review Process, NTP should improve the scientific rigor of the process.

As noted in its November 3, 2011 letter to Dr. Birnbaum, ACC is equally disappointed that NTP has provided a truncated and inadequate opportunity for stakeholder comment on NTP's proposed changes to the Review Process. ACC encourages NTP to implement the recommendations identified in ACC's November 3rd letter. A copy of that letter is also enclosed.

Given the global significance of the RoC, it is critical that NTP re-examine its Review Process periodically to ensure that it meets the highest standards of scientific integrity, transparency, and peer review. To assist that effort, ACC offers the enclosed comments on both the scientific methodology that NTP employs in preparing its RoC and the process by which peer review and public comments are sought, considered, and addressed.

ACC appreciates your thoughtful consideration of these comments and recommendations. In order to enhance the transparency and understanding of the RoC Review Process and NTP's perspective regarding it, ACC requests that NTP provide a written response to public comments it receives on its proposal. If you or NTP staff have any questions related to these comments, or would like additional information concerning them, please contact me by phone at 202-249-6405 or by e-mail at Rick_Becker@americanchemistry.com.

Sincerely,

[Redacted]

Richard A. Becker Ph.D., DABT
Senior Toxicologist
Regulatory and Technical Affairs Department

Enclosures

BEFORE THE
DEPARTMENT OF HEALTH AND HUMAN SERVICES

COMMENTS OF THE AMERICAN CHEMISTRY COUNCIL ON THE PROPOSED NATIONAL TOXICOLOGY PROGRAM (NTP) REVIEW PROCESS FOR THE REPORT ON CARCINOGENS

Proposed National Toxicology Program (NTP) Review Process for the Report on Carcinogens: Request for Public Comment and Listening Session 76 Fed. Reg. 67200 (Oct. 31, 2011)

November 30, 2011

EXECUTIVE SUMMARY

The American Chemistry Council ("ACC") appreciates the opportunity to comment on the National Toxicology Program's ("NTP") proposed revisions to the review process ("Review Process") for its Report on Carcinogens ("RoC"). 76 Fed. Reg. 67200 (Oct. 31, 2011). ACC has a longstanding interest in the RoC and its review process, and commented extensively on NTP's December 2003 proposed RoC review process.

Although NTP indicates that its proposed changes to the Review Process are intended to "*enhance* transparency and efficiency and to enable the NTP to publish the RoC in a timelier manner," NTP states that it merely wishes to "*maintain*" those elements of the existing process that go to "scientific rigor." 76 Fed. Reg. 67200 (emphasis added). The scientific shortcomings evident in the 12th RoC published this year demonstrate that improvements to *enhance* the scientific rigor of the RoC are also clearly needed.

For example, the National Academy of Sciences ("NAS") concluded in its independent review of the U.S. Environmental Protection Agency's ("EPA") draft Integrated Risk Information System ("IRIS") assessment of formaldehyde that EPA's review of the scientific literature as presented in the draft IRIS assessment does not provide a sufficient scientific basis for concluding that there is a causal link between formaldehyde exposure and leukemia. Even though EPA and NTP had reviewed and relied upon the same key studies, reports and underlying data in conducting their respective hazard evaluations, NTP determined in the 12th RoC that formaldehyde exposure is causally associated with leukemia. The differences in the conclusions of the NAS and the determinations in the 12th RoC regarding formaldehyde suggest that the Review Process lacks sufficient scientific robustness to ensure reliable assessments of carcinogenicity.

ACC recommends that NTP revise its proposed Review Process to ensure that the RoC meets the highest standards of scientific integrity, transparency, and peer review:

- **NTP should develop a sound scientific methodological approach to be utilized by NTP in preparing its RoC candidate substance evaluations.**
    - The approach currently employed by NTP lacks scientific rigor and fails to incorporate the fundamental analytical procedures critical to technically sound evaluations of the carcinogenic potential of candidate substances.
    - At a minimum, NTP should incorporate the recommendations in Chapter 7 of the NAS's Formaldehyde Report into the scientific evaluation of candidate substances.
    - If the methodological issues associated with the Review Process are not systematically addressed, future evaluations are likely to be flawed.

- **NTP should ensure meaningful involvement of stakeholders and the public.**

    - The truncated timeframes allotted for public comment in the proposed Review Process do not afford stakeholders sufficient time to effectively comment on the complex scientific issues addressed in the RoC. NTP should provide stakeholders with the time and information necessary to develop substantive comments.

    - There is no indication that submitted stakeholder comments are actually considered or addressed by peer reviewers or NTP. NTP should furnish a meaningful written response to public and partner agency comments.

    - The public should be given the opportunity to nominate experts to serve on an independent peer review committee, comment on the proposed members of the committee, and comment on the proposed charge questions to the committee.

- **NTP should incorporate peer review processes that foster substantive scientific dialogue among chemical-specific and subject matter-specific experts and promote opportunities for public comment.**

    - Rather than utilize the current NTP Board of Scientific Counselors, NTP should convene an independent peer review committee with the requisite technical expertise to evaluate the draft concept document and RoC Monograph for each candidate substance.

    - The peer review committee should not be limited to addressing narrow charge questions, but should provide all recommendations that the committee believes are germane to the scientific integrity of the RoC.

    - At a public meeting, commenters should be provided a reasonable opportunity to engage in a dialogue with the members of the peer review committee to allow for the thorough exploration of the various considerations relevant to the evaluation of candidate substances.

    - NTP should involve external chemical-specific and subject matter-specific experts at the early stages of drafting a RoC Monograph to enhance the quality of draft Monographs released for public and peer review.

    - NTP should issue a draft peer review report for public comment and convene a public meeting at which commenters can engage peer reviewers regarding concerns with the draft report.

ACC urges NTP to implement these recommendations to enhance the transparency and ensure the scientific quality of the RoC. These improvements will improve NTP decision-making regarding future RoCs and public confidence in their scientific integrity.

TABLE OF CONTENTS

INTRODUCTION

The American Chemistry Council ("ACC") appreciates the opportunity to comment on the National Toxicology Program's ("NTP") proposed revisions to the review process ("Review Process") for its Report on Carcinogens ("RoC"). 76 Fed. Reg. 67200 (Oct. 31, 2011). ACC has a longstanding interest in the RoC and its review process, and commented extensively on NTP's December 2003 proposed RoC review process.

The scientific, peer review, and public participation processes currently employed by NTP for the RoC are not adequate compared to processes that other federal agencies and scientific review bodies employ in hazard evaluations of potential carcinogens. In particular, the existing – and newly proposed – RoC process fails to foster substantive scientific dialogue among chemical-specific and subject matter-specific experts during the various stages of the process. The RoC – and the public's confidence in it – can undoubtedly benefit from a robust exchange of views among scientific experts from the public, academic, and private sectors.

Moreover, the RoC procedures for public/stakeholder participation fall well short of what is needed to ensure meaningful involvement of those parties in the RoC review process. Most notably, the current opportunities provided for public comment are inadequate and ineffective. The truncated timeframes allotted for public comment do not afford stakeholders sufficient time to effectively comment on the complex scientific issues addressed in the RoC. There is no indication that submitted stakeholder comments are actually considered or addressed by peer reviewers or NTP. To enhance the transparency of, and public involvement in, RoC preparation and decision-making, which are stated goals of NTP's proposed revisions to the RoC Review Process, NTP should remedy these important deficiencies.

Finally, NTP's "Proposed Report on Carcinogens Review Process" fails to address at all the scientific methodological approach to be utilized by NTP in preparing its RoC candidate substance evaluations. The approach currently employed by NTP lacks scientific rigor and fails to incorporate the fundamental analytical procedures critical to technically sound evaluations of the carcinogenic potential of candidate substances. Given the central role that the RoC plays in important public policy and private sector decisions regarding human health, it is imperative that the scientific processes employed for the RoC be improved substantially and in the manner reflected in these comments. The Review Process should reflect NTP's commitment to the highest standards of scientific inquiry.

I. NTP's Proposed Review Process Does Not Enhance the Scientific Rigor of the RoC

Although NTP indicates that its proposed changes to the Review Process are intended to "*enhance* transparency and efficiency and to enable the NTP to publish the RoC in a timelier manner," NTP states that it merely wishes to "*maintain*" those elements of the existing process that go to "scientific rigor." 76 Fed. Reg. 67200 (emphasis added). The scientific shortcomings evident in the 12th RoC published this year demonstrate that improvements to *enhance* the scientific rigor of the RoC are also clearly needed.

The remarkable differences in the conclusions of the National Academy of Sciences panel ("NAS") in its peer review report regarding the U.S. Environmental Protection Agency's ("EPA") draft Integrated Risk Information System ("IRIS") assessment of formaldehyde[1] and the NTP determinations in the 12th RoC regarding formaldehyde suggest that the Review Process lacks sufficient scientific robustness to ensure reliable assessments of carcinogenicity. NAS was charged with conducting an independent scientific review of the EPA's draft IRIS assessment of formaldehyde. EPA's IRIS assessment contains a hazard analysis, very similar to NTP's substance profile for formaldehyde.

NAS significantly challenged EPA's evaluation of the possible relationship between formaldehyde exposure and human leukemia. The committee concluded that EPA's claims that formaldehyde causes leukemia, myeloid leukemia or related hematopoietic cancers are not supported in EPA's assessment. The committee noted that EPA's preliminary conclusion that the data supported a causal relationship appeared to be "subjective" in nature, and that no clear scientific framework had been applied by EPA in reaching that conclusion. The absence of such a framework was judged by the committee as troublesome, given that the available scientific evidence is very weak. As emphasized by NAS, "the absence of a causal framework for these cancers is particularly problematic given the inconsistencies in the epidemiologic data, the weak animal data, and the lack of mechanistic data."[2] Moreover, NAS stated that the epidemiologic data were limited by "uncertainties of exposure assessment, possible confounding by other pollutants, and reliance on mortality data rather than incidence data…."[3]

NAS's opinion was that EPA's review of the scientific literature as presented in the draft IRIS assessment does not provide a sufficient scientific basis for concluding that there is a causal link between formaldehyde exposure and leukemia. EPA and NTP had reviewed and relied upon the same key studies, reports and underlying data in conducting their respective hazard evaluations of the possible relationship between formaldehyde exposure and leukemia and other lymphohematopoietic malignancies. NAS's findings support the conclusion that the determination in the 12th RoC that formaldehyde exposure is causally associated with leukemia – seemingly regardless of the level of exposure – is not supported by best available science.

II. NTP Should Incorporate the NAS's Recommendations into the Scientific Evaluation of Candidate Substances

In its April 2011 Formaldehyde Report, NAS also devoted an entire chapter to addressing persistent and pervasive shortcomings in the manner in which scientific information is assessed, weighed, and presented under IRIS. Although the same systemic deficiencies identified by NAS pervade NTP's RoC program, inexplicably NTP and the Department of Health and Human Services ("HHS") dismissed the NAS Formaldehyde Report as being of "limited applicability" to the 12th RoC.

---

[1] See "Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde, National Research Council of the National Academies, April 2011) ("NAS Formaldehyde Report").

[2] *Id.* at 8.

[3] *Id.* at 83.

ACC concurs with NAS's assessment that if the methodologic issues associated with IRIS assessments are not systematically addressed, future assessments are likely to have the same unfortunate – but avoidable – problems. Similarly, future RoCs likely will be flawed, unless NTP incorporates the recommendations set forth by NAS in Chapter 7 of its Formaldehyde Report into its Review Process.[4] Because NAS considers those recommendations to be "critical for the development of a scientifically sound assessment,"[5] NTP should incorporate those recommendations as part of the Review Process and commit to implement them in preparing future RoCs. A copy of Chapter 7 of the NAS Formaldehyde Report is appended to these comments and incorporated by reference.

In particular, but without intending to limit the scope and application of the NAS recommendations, NTP should expressly incorporate into the Review Process the following elements for evaluating candidate substances:

- All evaluations should implement the NAS recommendations for reviewing studies, assessing weight of evidence, determining mode of action, and establishing cause and effect.

- All evaluations should rely on best available scientific information regarding hazard and exposure, employ consistent and objective "state of the science" methods and models, and utilize transparent data evaluation procedures for evaluating both data quality and for weighing the full body of scientific evidence.

- NTP should consistently follow Office of Management and Budget ("OMB"), HHS, and National Institute of Health ("NIH") information quality and peer review guidelines, as applicable, for all evaluations.

- In all evaluations, NTP should employ proven methods for establishing cause and effect and, where it exists, reliable data in lieu of default assumptions.

- To ensure transparency of its evaluations, NTP should provide full disclosure of the following:

  – data, methods, and models sufficient to allow independent reanalysis by qualified persons;

  – the rationale for choosing methods and models;

  – a template or flow diagram illustrating requirements of applicable federal guidance and explaining any instances in which that guidance is not followed and why;

[4] ACC recognizes that the purpose and scope of the RoC are not identical to those of EPA IRIS assessments. Nonetheless, the nature of the scientific evaluation undertaken is sufficiently similar that the NAS recommendations are equally germane for the RoC.

[5] NAS Formaldehyde Report, at 121.

- assumptions and extrapolations used and their impact on the evaluation;
- the impact on the evaluation of use of models vs. measurements;
- plausible alternative modes of action, models and assumptions, the rationale for choices made among those alternatives, and the impacts of one choice vs. another on the evaluation;
- significant knowledge/data gaps and other sources of uncertainty, and their implications for the evaluation;
- identification of where default assumptions or policy decisions, rather than conclusions based on scientific evidence, are involved;
- the relative strength of each component of the evaluation and its impact on the strength of the overall evaluation; and
- major conclusions and NTP's degree of confidence in them.

- NTP should employ a consistent weight of evidence framework, formulated upon a hypothesis-based mode of action evaluation procedure, so that data from all relevant studies can be systematically reviewed, given appropriate weight, and integrated in a manner that provides a robust understanding of the mode of action. "All relevant studies" may include, for example, guideline toxicity studies, studies published in the peer-reviewed literature, Good Laboratory Practices ("GLP") and non-GLP investigations, studies demonstrating positive or negative findings, epidemiological investigations, specialized mode of action studies, and the like.

These principles would implement the recommendations in Chapter 7 of the NAS Formaldehyde Report in the RoC Review Process and significantly enhance the transparency of the RoC. They would also add scientific rigor that is critical to ensuring policymaker and public confidence in RoC evaluations that are employed in important public policy and private sector decisions.

III. NTP's Review Process Should Provide for Rigorous Public and Peer Review of RoC Monographs

A. Peer Review of RoC Monographs Must Comply with OMB and HHS Information Quality Requirements

All peer review of draft RoC Monographs should comply with the requirements for peer review set forth in the OMB "*Final Information Quality Bulletin for Peer Review*" (December 16, 2004) ("OMB Peer Review Guidelines"). The OMB Peer Review Guidelines establish government-wide guidance "aimed at enhancing the practice of peer review of government science documents" and thereby "improv[ing] the quality of scientific information

upon which policy decisions are based." Among other things, OMB's Peer Review Guidelines provide guidance on what information is subject to peer review, the selection of appropriate peer reviewers, peer review planning processes, opportunities for public participation, and related issues. Because the RoC constitutes a "highly influential scientific assessment" as defined in the OMB Peer Review Guidelines, peer review should meet the OMB's requirements for such assessments.[6]

In its final Information Quality Act ("IQA") Guidelines ("OMB IQA Guidelines"), OMB emphasized the importance of "competent and credible peer review" by federal agencies and, for transparency, documentation of compliance with criteria regarding (i) the selection of peer reviewers with the range of necessary technical expertise, (ii) disclosure of prior technical/policy positions peer reviewers may have taken on the issues at hand, (iii) disclosure of sources of personal or institutional funding (private or public sector), and (iv) an open and rigorous peer review process.[7] NTP is required to comply with these, and other, aspects of the OMB IQA Guidelines.[8] The final revised Review Process should therefore specifically state that the process will comply with the OMB Peer Review and IQA Guidelines.

In its own IQA guidelines, NIH addresses preparation of RoCs to ensure that they meet the "objectivity" and "utility" requirements of the OMB IQA Guidelines. In those guidelines, NIH recognizes that RoCs constitute "influential scientific information" subject to the heightened IQA objectivity standards for such information.[9] Inasmuch as RoCs certainly have a clear and substantial impact on important public policies and/or important private sector

---

[6] See OMB Peer Review Guidelines, at 23 (defining "highly influential scientific assessments" to include, among other things, assessments that have "significant interagency interest," which is clearly the case for the RoC given the number, size, and range of "agency partners" NTP identifies in footnote 2 of its proposed Review Process).

[7] See 67 Fed. Reg. 8452, 8454, 8459 (Feb. 22, 2002). Ensuring that peer review panels are composed of highly qualified experts across the range of necessary disciplines is essential to competent and credible peer review and should be the principal criterion for selection of peer reviewers. Past and present affiliations and/or funding sources that do not represent clear conflicts of interest should not disqualify potential peer review candidates with necessary expertise. To address potential conflict issues, policies and procedures for full disclosure of conflicts of interest and bias, as described in the NAS Conflict of Interest Policy, should be implemented. Finally, it is important that there be a balance of scientific perspectives among the peer review panel and that peer reviewers devote sufficient time to a rigorous peer review. Moreover, to enable a thorough peer review, the NTP should ensure a high degree of transparency about data and methods to facilitate the reproducibility of such information by the peer review panel. The high degree of transparency specifically applies to (1) the source of the data used, (2) the various assumptions employed, (3) the analytic methods, and (4) the statistical procedures employed.

[8] See, e.g., 67 Fed. Reg. 8452 ("[i]t is crucial that information Federal agencies disseminate meets these guidelines); *Prime Time Int'l Co. v. Vilsack, et al.*, 599 F.3d 678, 685 (D.C. Cir. 2010) (OMB IQA Guidelines are "binding" on federal agencies); NIH IQA Guidelines, § V.1 (NIH "must adhere to the … OMB Information Quality Guidelines").

[9] See NIH IQA Guidelines, §§ V.2.d, fourth paragraph; VII (identifying NIH research reports as potentially being considered "influential" information).

decisions, ACC concurs that RoCs are "influential scientific information" subject to those heightened standards.[10]

Accordingly, the revised Review Process should expressly state that that process will comply in all respects with the OMB and NIH IQA Guidelines applicable to influential scientific information. The NIH IQA Guidelines, § VII, set forth requirements applicable to influential scientific information, including (i) the obligation to state clearly how analytic results are generated by making transparent "the specific data used, various assumptions, specific analytic methods, statistical procedures, [and] sources of error," (ii) the requirement that scientific conclusions be based on best available science and supporting studies, particularly peer-reviewed studies, conducted in accordance with sound and objective scientific practices," and (iii) a commitment to "rigorous peer review" designed to obtain input from "qualified reviewers . . . for accuracy, completeness, and quality."

The importance of an express commitment to applicable IQA guidelines is underscored, for example, by the IQA Request for Correction ("RFC") submitted on October 26, 2009 by the Styrene Information and Research Center regarding the "Final Report on Carcinogens Background Document for Styrene," issued by NTP on September 29, 2008. That RFC documents a large number of fundamental IQA deficiencies associated with the styrene background document prepared for purposes of the 12th RoC. An upfront commitment to conform RoC information disseminated by NTP to applicable IQA requirements, including those related to pre-dissemination review processes and standards, should minimize the time and resources NTP will need to devote to IQA challenges to that information.

B. NTP Should Propose Specific Measures to Improve Public and Peer Review During the Review Process

Although NTP states that it seeks to "enhance" the transparency and efficiency of the Review Process and "maintain" elements of it related to public involvement and external peer review, the October 31 Federal Register notice does not identify the specific proposed changes to the review process that was employed in the 12th RoC to achieve those objectives. Nor does NTP explain why it believes that the modifications it is proposing will achieve those goals.

The proposed Review Process appears to provide less explanation of the various steps of the process than the review process description for the 12th RoC. Consequently, it is unclear whether NTP is merely providing a more succinct summary of the Review Process or, for certain elements of the process, actually proposing substantive changes or deletion of elements of the process. For example, the description of the process for the 12th RoC provided an explanation of the types of information gathered for discussion in what NTP now refers to as the "cancer evaluation component" of the draft RoC Monograph. In contrast, the proposed Review Process merely identifies those topics without describing their content. In order to enhance the transparency of the Review Process, the description of it should explicitly and comprehensively set forth the actual steps, information, and scientific methodology contemplated.

[10] See OMB IQA Guidelines, § V.9; NIH IQA Guidelines, § VII (definition of "influential," including – for purposes of NIH – information that "will have important consequences for specific health practices").

Because the scope of NTP's proposed changes and of the Review Process are not clear, ACC sets forth below elements of a high quality review process that are not addressed specifically or adequately in NTP's proposed Review Process. These comments are intended to ensure that the Review Process meets the Obama Administration's stated commitment to transparency and effective public participation and peer review in scientific assessments. ACC's comments are presented separately for each of the four parts of the proposed Review Process: (i) nomination and selection of candidate substances, (ii) scientific evaluation of candidate substances, (iii) public release of the draft RoC Monograph and peer review, and (iv) HHS approval and release of the latest edition of the RoC.

1. Nomination and Selection of Candidate Substances

To provide a meaningful opportunity for public comment on the proposed nomination and selection of candidate substances and effective peer review, NTP should augment its proposed process so that it explicitly incorporates the following elements:

- Upon the deadline for nominations of substances, NTP should promptly establish a docket that makes all documentation received regarding such substances available electronically to the public. All such submissions, including those from NTP's "agency partners," should also be available for copying upon submission. This step would maximize the time that submissions are available for public evaluation.

- For purposes of public comment on the draft concept documents for each substance proposed for evaluation, NTP should establish – before the start of the public comment period – a docket that includes all public and agency submissions, the draft concept documents for each substance proposed for evaluation, and all supporting materials used by NTP to develop its draft concept documents. The docket, and index to the docket, should be promptly updated as additional submissions from commenters, including NTP's agency partners, are received.

- The period for public comment on the draft concept document for each substance proposed for evaluation should be at least 90 days.

- Rather than utilize the current NTP Board of Scientific Counselors ("BSC"), NTP should convene an independent peer review committee with the requisite technical expertise ("Peer Review Committee"), to evaluate the draft concept document for each substance proposed for evaluation, and during a public meeting, provide for the following:

  – NTP should provide the public with an adequate opportunity to comment on proposed charge questions, before they are provided to the Peer Review Committee. While NTP may specifically request that the Peer Review Committee provide comment on select aspects of the draft concept documents, the Peer Review Committee should be instructed that it is free to comment on any issue relevant to the nomination of candidate substances, including, without limitation, whether a substance proposed

for evaluation in the RoC should be selected for inclusion in the RoC. If NTP rejects charge questions proposed by public commenters, it is imperative that NTP provide a written explanation of why it has done so.

- The public should be provided access to the full docket for the RoC report at least 90 days before the public meeting. This time period would give interested persons a reasonable opportunity to review relevant materials, thereby affording a meaningful opportunity to present comments at that meeting.

- In order to ensure a sufficient opportunity to convey scientific oral remarks at the public meeting, each presenter should be provided up to 30 minutes. Depending upon the number of registered commenters, this time allotment may necessitate a longer public meeting than originally anticipated. The opportunity for meaningful oral comments should not be abridged simply because of a large number of presenters.

- At the public meeting, presenters should be provided a reasonable opportunity to engage in a dialogue with the Peer Review Committee, either as part of the presenter's allotted time or in a separate session once all commenters have presented their oral comments. This exchange of views (which should encompass questions from and to the Peer Review Committee to enhance transparency and understanding) is critical to an exploration of the various considerations relevant to the selection of substances for evaluation in the RoC.

- This step of the Review Process involves not only a determination of which substances are to be selected for evaluation in the RoC but also (i) information on exposure and extent of evidence of carcinogenicity, and (ii) the proposed approach for development of the cancer evaluation component of the draft RoC Monograph. Given the substantive nature of these issues and the consequences of the determinations, for scientific integrity and transparency, NTP needs to furnish a meaningful written response to public or partner agency comments it receives on draft concept documents. ACC notes that NTP posted a response to certain public comments it received in preparing the 12th RoC on its website. The responses, however, were not a meaningful effort to fully address the issues being raised by the commenters.

2. Scientific Evaluation of Candidate Substances

Although the nature of the scientific evaluation of candidate substances is addressed earlier in these comments, certain "process" issues related to public and peer review of that evaluation are also important. First and foremost, it is critical that *all* information on exposure and properties of candidate substances that are employed in preparation of draft RoC Monographs be placed in the docket and made available to the public and peer reviewers before peer and public review and comment commence.

The proposed Review Process states that NTP's approach to evaluating the scientific information on a candidate substance for purposes of developing the cancer evaluation component of a draft RoC Monograph

> is tailored to enable [the Office of the Report on Carcinogens ("ORoC")] to use the most appropriate mechanism(s) to obtain external advice and address scientific issues for assessing the carcinogenicity of a given candidate substance *and may vary among substances*. The approach *may* include external scientific input (e.g., expert panel, workshop, individual technical advisors), public input (e.g., listening sessions, comment), *and/or* interagency input. (emphasis added)

This statement suggests that some "cancer evaluations" will receive the substantial benefit of external scientific, public, and interagency input but that others will not (or will receive much more limited input). There is no reason why that should be the case, particularly where NTP has not identified any criteria that would guide discretionary determinations as to whether that input will be sought and, if so, from whom, in what manner, and to what extent. The proposed Review Process should be revised to provide for a uniform, minimum opportunity for external scientific, public, and interagency input (as set forth in these comments) in the development of all "cancer evaluations" before the preparation of draft RoC Monographs. Early input from these sources will enhance the scientific integrity and transparency of the development of draft Monographs. Moreover, input at this stage may actually accelerate the development of RoC Monographs by ensuring that key issues can be brought to NTP's attention early in the process. NTP should also expand the list of "partner agencies" to include OMB, NASA, and all other agencies that are typically part of interagency review but are not on the NTP Executive Committee.

In particular, the involvement of external chemical-specific and subject matter-specific experts at the early stages of drafting a RoC Monograph would enhance the quality of draft Monographs released for public and peer review.

Furthermore, NTP proposes that the BSC comment on "the proposed approach for obtaining external scientific and public inputs in developing the cancer evaluation component of its draft RoC Monograph." NTP should initiate a full review, including solicitation of public comments, on the approaches NTP should adopt in developing the cancer evaluation component of draft RoC Monographs. Moreover, NTP should immediately seek NAS input on 1) the criteria that NTP should utilize in selecting a candidate substance for evaluation, and 2) NTP's current listing/delisting approach for evaluating studies, integrating results using weight of evidence and the criteria for classification of substances as "known" or "reasonably anticipated" carcinogens.

3. Public Release of Draft RoC Monograph and Peer Review

To ensure meaningful public comment on draft RoC Monographs the review process should incorporate the following elements:

- All documentation evaluated by NTP in preparation of a draft Monograph should be made available for public review and comment no later than the start of the public comment period.

- The period for public comment on draft Monographs should be no less than 90 days.

- The public should be given access to all documentation evaluated by NTP in preparing the draft Monograph at least 30 days before the meeting of the "external advisory group" to peer review the revised draft Monograph. ACC notes that the process for the 12th RoC provided that NTP would publish a Federal Register notice announcing the public meeting regarding, and availability of, draft "background documents" and "substance profiles" at least 60 days prior to the expert panel meeting on those documents. NTP's new proposed process is silent as to that time period.

- Each presenter at the public meeting should be provided up to 30 minutes for comments and, as set forth above, should be provided a reasonable opportunity to engage in a meaningful exchange of views with the peer reviewers and NTP to enhance the value of public input (which exchange should encompass questions from or to peer reviewers and NTP to augment transparency and increase understanding).

NTP's proposal states that NTP will convene a meeting of "an external advisory group (e.g., BSC or expert panel) to review the revised draft RoC Monograph." The composition of the group of scientists used to peer review draft Monographs is a critical element of competent and credible peer review and should not be addressed in such a cursory fashion in the Review Process description.[11] NTP should create separate expert panels with requisite expertise to peer review the draft RoC monographs. Moreover, the public must have the opportunity to nominate experts with relevant scientific disciplines to serve on the expert peer review panel. The process NTP established for the 12th RoC expressly provided for that opportunity. Without explanation, the Review Process now proposed by NTP is silent on the subject, suggesting that NTP no longer intends to afford that opportunity. NTP needs to state explicitly that that opportunity – and a meaningful opportunity to address the peer review considerations discussed below – will be provided.

As noted above, the proposed Review Process does not provide for the public to nominate scientists for the peer review conducted at either the selection of candidate substances stage or at the peer review of draft RoC Monograph stage. Nor does it indicate who will comprise the "external advisory group" providing the peer review, other than observing that the "BSC or expert panel" are *examples* of who might provide peer review. In order to ensure that the OMB guidelines and other principles relevant to transparent, competent, and credible peer review are adhered to, it is critical that the revised Review Process specifically provide for an early and meaningful opportunity for the public to nominate scientists to serve on the peer review panel. It is also important that NTP disclose the peer reviewers it proposes to use; disclose relevant prior technical/policy positions, funding sources, and potential conflicts of

[11] For example, in footnote 6 of its Review Process description for the 12th RoC, NTP acknowledged the need for an expert panel with "relevant expertise and knowledge selected by NTP from the public and private sectors." NTP also indicated that a peer review panel should consist of a "balanced and unbiased group of highly qualified individuals" selected "in accordance with the Federal Advisory Committee Act and HHS implementing regulations." NTP's proposed revised process description makes no mention of these core principles.

interest of proposed peer reviewers; and provide a meaningful opportunity for the public to comment on that information before a final peer review panel is established.

According to the proposed revised Review Process, the charge to peer reviewers "*is twofold*: (1) to comment on the cancer evaluation component, specifically whether it is technically correct and clearly stated, whether the NTP has objectively presented and assessed the scientific evidence, and whether the scientific evidence is adequate for applying the listing criteria, and (2) to comment on the substance profile, specifically, whether the scientific justification presented in the substance profile supports the NTP's preliminary decision on the listing status of the candidate substance in the RoC." ACC concurs that these two charges are appropriate. However, the charge should not be so limited.

First, the public should be provided a meaningful opportunity to recommend specific charges. NTP should commit to thoughtful consideration of these public recommendations, and incorporate these into the final set of charge questions provided to the peer reviewers. In instances where NTP decides not to pose a recommended relevant charge question, NTP should provide a clear written justification that is scientifically sound and transparent.

Second, the peer review panel should be able to – and expressly encouraged to – bring to the attention of NTP any recommendation or consideration that the panel believes is germane to the scientific integrity of the RoC. These recommendations and considerations could include, for example, whether NTP adequately (i) justified reliance on, or rejection of, certain studies, (ii) discussed the strengths and weaknesses of studies relied on, (iii) conducted a weight of evidence evaluation involving the studies, (iv) addressed all available scientific information in a comprehensive fashion, and (v) evaluated uncertainties associated with the scientific information assessed. Peer reviewers may wish to comment as well on data analysis, mode of action, human relevance and extrapolation method considerations. Finally, the peer review panel may also want to recommend or encourage NTP to address certain specific public comments on a draft Monograph.

The proposed Revised Review process fails to provide for issuance of a draft peer review report for public comment or a meeting at which public commenters can engage peer reviewers regarding concerns with that draft report, so that those reviewers have the benefit of the views of other outside scientific experts on complex scientific issues. The Review Process should be revised to provide for:

- A period of at least 90 days for public review of a draft peer review report before a public meeting.

- A public meeting at which each commenter has up to 30 minutes to present oral comments.

- A meaningful opportunity for commenters at that meeting to engage in an exchange of views and questions regarding pertinent scientific information.

- A public comment period of at least 30 days after the conclusion of the peer review meeting during which the public may submit written comments to peer reviewers. Peer reviewers will be instructed by NTP to carefully review and

evaluate all germane public comments prior to issuing their final peer review report.

- In addition to requiring that NTP prepare a response to the final peer report that is released to the public, NTP should prepare a written response to all substantive public comments.[12] For efficiency, similar or related comments may be aggregated or combined by NTP. At a minimum, the response to comments should be provided regardless of whether NTP agrees with the public comment involved (if it does not, NTP should explain why), and identify how and why the draft report was modified to account for public or peer review comments.

The Review Process for the 12th RoC provided for internal government reviews following the expert panel meeting regarding draft background documents. That internal review process, consisting of separate meetings of an interagency scientific review group ("ISRG") and the NIEHS/NTP scientific review group ("NSRG"), was designed to solicit the independent recommendations of those groups on the candidate substances and what the listing status of those substances should be. For purposes of that review, the ISRG and NSRG were to be provided all relevant information, including the expert panel report and all public comments submitted by that stage of the process. Without explanation, NTP has eliminated these governmental reviews from the proposed revised Review Process. A timely and formal interagency review and comment period would be useful subsequent to the one expressly provided for by NTP at the nomination/selection of candidate substance stage.

4. HHS Approval and Release of Latest Edition of RoC

The revised Review Process should state that HHS preparation, review, approval, and release of each new edition of the RoC will be in accordance with NIH Manual Chapter 1183: NIH Publications and Audiovisuals: Preparation, Review, Approval, and Distribution, January 1, 2009. Among other things, that chapter of the NIH Manual seeks to ensure compliance with the OMB and NIH IQA Guidelines and OMB Peer Review Guidelines.

Unlike the proposed revised Review Process, the process employed for the 12th RoC stated that after NTP responded to the peer review report, the "final draft" RoC was to be submitted to the NTP Executive Committee not only for "consultation," but also for "review and comment." The final Review Process should expressly provide that the peer review report, and NTP's response to that report and to public comment, are to be furnished to the NTP Executive Committee and NTP's agency partners for their review in commenting upon the final draft RoC. In addition, the revised process should provide for posting of the NTP Executive Committee's comments on the NTP website (or docket) when the final RoC is published.

[12] As noted above, NTP committed to preparing a response to public comments in connection with the 12th RoC and indicated that it would "assess the merits of responding to public comments following completion" of that RoC. In its October 31 Federal Register notice, NTP provided no rationale as to why a response to public comments was not warranted. NTP's stated objectives of enhancing the transparency of the Review Process and maintaining the "critical element" of public involvement can only be met if a meaningful and timely response to public comments is provided. That response is necessary both to ensure that public comments have been considered in good faith and to instill confidence that public input is in fact a critical component of the Review Process.

When NTP issues its response to the peer review report, it should also issue a written response to all substantive public comments, as noted above. Doing so will both ensure that public comments, including those of outside scientific experts with knowledgeable insights, have been given due consideration, and enhance transparency and public understanding of the basis for NTP's conclusions. NTP's review process for the 12th RoC recognized the importance of issuing publicly a written response to comments.

IV. Conclusions

ACC commends NTP for seeking public input on how the RoC Review Process can be improved so as to enhance its transparency and ensure its scientific quality. Unfortunately, NTP's proposals result in a missed opportunity to institute methods by which the scientific rigor of the RoC can be improved. ACC is also concerned that the proposed Review Process limits the opportunity for, and effectiveness of, peer review and public comment that could vastly improve the quality of RoCs. Implementation of ACC's recommendations herein would remedy both of those shortcomings and substantially advance both NTP decision-making regarding future RoCs and public confidence in their scientific integrity.

++

Attachments:

- ACC Correspondence to the Director of NTP dated November 5, 2011
- Chapter 7 "Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde." NAS, 2011.

CAL DOOLEY
PRESIDENT AND CEO

November 3, 2011

Linda S. Birnbaum, Ph.D., D.A.B.T., A.T.S.
Director, NIEHS & NTP
National Institute of Environmental Health Sciences
P.O. Box 12233
Mail Drop B2-01
Research Triangle Park, North Carolina 27709
Tel (919) 541-3201
Fax (919) 541-2260
birnbaumls@niehs.nih.gov

Re: Proposed National Toxicology Program (NTP) Review Process for the Report on Carcinogens: Request for Public Comment and Listening Session

Dear Dr. Birnbaum:

We are pleased that the National Toxicology Program (NTP) recognizes the need for improvements in the Report on Carcinogens (RoC), and appreciate NTP's interest in obtaining stakeholder input on the "Proposed Report on Carcinogens Review Process" dated October 31, 2011.[1,2] All stakeholders expect the RoC to be firmly based on up-to-date scientific knowledge, meet the highest of standards of scientific inquiry, and be evaluated in accordance with acceptable scientific approaches. Unfortunately, the current policies, practices and resulting assessments of the RoC do not consistently meet these standards, and improvements are clearly warranted.

ACC is concerned that timeframe and structure presented for NTP's process for soliciting stakeholder perspectives on improvements to the RoC will not produce meaningful and substantive input. We are also concerned that the process outlined for public input is not adequate given the substantive scientific nature of the issues raised by the RoC. We are, therefore, writing to request that both the timeline and process be revised as specifically outlined later in this letter. In light of the truncated comment period, we would appreciate a response to our request for an extension of the comment period as soon as possible.

On October 31, 2011, NTP published its request for public comment on revisions to the RoC process, allowing only three weeks for the public to submit oral statements and/or slides for the public listening session by November 21, 2011. Moreover, registration to present oral remarks is unnecessarily limited to the first 15 registrants. Written comments are due on November 30, just

[1] Federal Register / Vol. 76, No. 210 / Monday, October 31, 2011, page 67200
[2] http://ntp.niehs.nih.gov/NTP/RoC/Thirteenth/Process/ProposedROCReviewProcess2011.pdf

days after the Thanksgiving holiday. In sum, NTP's timelines are not adequate to obtain, review and respond to meaningful public input. In ACC's view, this approach is not consistent with President Obama's goals for transparency and scientific integrity.

The President pledged that the "science and scientific process must inform and guide decisions of my Administration." A Presidential Memorandum directs each agency to use well-established scientific processes to inform public policy decisions. As you are aware, many questions have arisen regarding the scientific evaluation and peer review processes used to develop the RoC. The striking differences between the National Research Council's (NRC) peer review report of the draft Integrated Risk Information System assessment and the NTP findings regarding the causal relationship between formaldehyde exposure and leukemia is but one example of the need for NTP to substantively improve the RoC process. NTP must provide the public with a meaningful opportunity to comment on NTP's proposed RoC process changes.

Therefore, ACC requests that NTP take the following actions:

1) Extend the written comment period to 90 days;
2) Replace the web-only format proposed by NTP for receiving oral public input with an in-person public meeting that is simultaneously webcast;
3) Reschedule the public meeting to two to three weeks after the submission deadline for written comments;
4) Structure the public meeting as a dialogue, in which NTP staff actively engage in discussions with stakeholders on the substantive issues, rather than as a one-way discourse; and
5) Following the public meeting, NTP should analyze and respond to public comments, providing a written record of NTP's rationale for accepting or rejecting comments and for making specified policy choices, and publish a revised draft process clearly noting what changes have been made.

ACC also recommends that, after the actions above are completed, NTP submit the revised draft RoC process, including the improved scientific analysis procedures, for independent scientific peer review by the NRC. This will assure that before the next RoC is developed, the scientific procedures used in the RoC meet the highest standards of scientific inquiry.

Sincerely,

[Redacted]

Cal Dooley
President and CEO

cc: HHS Secretary Kathleen Sebelius

Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde

Committee to Review EPAâ€™s Draft IRIS Assessment of Formaldehyde; National Research Council

ISBN: 0-309-21194-8, 194 pages, 6 x 9, (2011)

7

A Roadmap for Revision

In reviewing the draft assessment *Toxicological Review of Formaldehyde-Inhalation Assessment: In Support of Summary Information on the Integrated Risk Information System (IRIS)*, the committee initially evaluated the general methodology (Chapter 2) and then considered the dosimetry and toxicology of formaldehyde (Chapter 3) and the review of the evidence and selection of studies related to noncancer and cancer outcomes (Chapters 4 and 5). Finally, the committee addressed the calculation of the reference concentrations (RfCs) for noncancer effects and the unit risks for cancer and the treatment of uncertainty and variability (Chapter 6). In this chapter, the committee provides general recommendations for changes that are needed to bring the draft to closure. On the basis of "lessons learned" from the formaldehyde assessment, the committee offers some suggestions for improvements in the IRIS development process that might help the Environmental Protection Agency (EPA) if it decides to modify the process. As noted in Chapter 2, the committee distinguishes between the process used to generate the draft IRIS assessment (that is, the development process) and the overall process that includes the multiple layers of review. The committee is focused on the development of the draft IRIS assessment.

CRITICAL REVISIONS OF THE CURRENT DRAFT IRIS ASSESSMENT OF FORMALDEHYDE

The formaldehyde draft IRIS assessment has been under development for more than a decade (see Chapter 1, Figure 1-3), and its completion is awaited by diverse stakeholders. Here, the committee offers general recommendations—in addition to its specific recommendations in Chapters 3-6—for the revisions that are most critical for bringing the document to closure. Although the committee suggests addressing some of the fundamental aspects of the approach to generating the draft assessment later in this chapter, it is not recommending that the assessment for formaldehyde await the possible development of a revised approach. The following recommendations are viewed as critical overall changes needed to complete the draft IRIS assessment:

- To enhance the clarity of the document, the draft IRIS assessment needs rigorous editing to reduce the volume of text substantially and address redundancy and inconsistency. Long descriptions of particular studies, for example, should be replaced with informative evidence tables. When study details are appropriate, they could be provided in appendixes.
- Chapter 1 needs to be expanded to describe more fully the methods of the assessment, including a description of search strategies used to identify studies with the exclusion and inclusion criteria clearly articulated and a better description of the outcomes of the searches (a model for displaying the results of literature searches is provided later in this chapter) and clear descriptions of the weight-of-evidence approaches used for the various noncancer outcomes. The committee emphasizes that it is not recommending the addition of long descriptions of EPA guidelines to the introduction, but rather clear

concise statements of criteria used to exclude, include, and advance studies for derivation of the RfCs and unit risk estimates.

- Standardized evidence tables for all health outcomes need to be developed. If there were appropriate tables, long text descriptions of studies could be moved to an appendix or deleted.
- All critical studies need to be thoroughly evaluated with standardized approaches that are clearly formulated and based on the type of research, for example, observational epidemiologic or animal bioassays. The findings of the reviews might be presented in tables to ensure transparency. The present chapter provides general guidance on approaches to reviewing the critical types of evidence.
- The rationales for the selection of the studies that are advanced for consideration in calculating the RfCs and unit risks need to be expanded. All candidate RfCs should be evaluated together with the aid of graphic displays that incorporate selected information on attributes relevant to the database.
- Strengthened, more integrative, and more transparent discussions of weight of evidence are needed. The discussions would benefit from more rigorous and systematic coverage of the various determinants of weight of evidence, such as consistency.

FUTURE ASSESSMENTS AND THE IRIS PROCESS

This committee's review of the draft IRIS assessment of formaldehyde identified both specific and general limitations of the document that need to be addressed through revision. The persistence of limitations of the IRIS assessment methods and reports is of concern, particularly in light of the continued evolution of risk-assessment methods and the growing societal and legislative pressure to evaluate many more chemicals in an expedient manner. Multiple groups have recently voiced suggestions for improving the process. The seminal "Red Book," the National Research Council (NRC) report *Risk Assessment in the Federal Government: Managing the Process*, was published in 1983 (NRC 1983). That report provided the still-used four-element framework for risk assessment: hazard identification, dose-response assessment, exposure assessment, and risk characterization. Most recently, in the "Silver Book," *Science and Decisions: Advancing Risk Assessment*, an NRC committee extended the framework of the Red Book in an effort to make risk assessments more useful for decision-making (NRC 2009). Those and other reports have consistently highlighted the necessity for comprehensive assessment of evidence and characterization of uncertainty and variability, and the Silver Book emphasizes assessment of uncertainty and variability appropriate to the decision to be made.

Science and Decisions: Advancing Risk Assessment made several recommendations directly relevant to developing IRIS assessments, including the draft formaldehyde assessment. First, it called for the development of guidance related to the handling of uncertainty and variability, that is, clear definitions and methods. Second, it urged a unified dose-response assessment framework for chemicals that would link understanding of disease processes, modes of action, and human heterogeneity among cancer and noncancer outcomes. Thus, it suggested an expansion of cancer dose-response assessments to reflect variability and uncertainty more fully and for noncancer dose-response assessments to reflect analysis of the probability of adverse responses at particular exposures. Although that is an ambitious undertaking, steps toward a unifying framework would benefit future IRIS assessments. Third, the Silver Book recommended that EPA assess its capacity for risk assessment and take steps to ensure that it is able to carry out its challenging risk-assessment agenda. For some IRIS assessments, EPA appears to have difficulty in assembling the needed multidisciplinary teams.

The committee recognizes that EPA has initiated a plan to revise the overall IRIS process and issued a memorandum that provided a brief description of the steps (EPA 2009a). Figure 7-1 illustrates the steps outlined in that memorandum. The committee is concerned that little information is provided on what it sees as the most critical step, that is, completion of a draft IRIS assessment. In the flow diagram, six steps are devoted to the review process, and thus the focus of the revision appears to be on the steps

after the assessment has been generated. Although EPA may be revising its approaches for completing the draft assessment (Step 1 in Figure 7-1), the committee could not locate any other information on the revision of the IRIS process. Therefore, the committee offers some suggestions on the development process.

In providing guidance on revisions of the IRIS development process (that is, Step 1 as illustrated in Figure 7-1), the committee begins with a discussion of the current state of science regarding reviews of evidence and cites several examples that provide potential models for IRIS assessments. The committee also describes the approach now followed in reviewing and synthesizing evidence related to the National Ambient Air Quality Standards (NAAQSs), a process that has been modified over the last 2 years. It is provided as an informative example of how the agency was able to revise an entrenched process in a relatively short time, not as an example of a specific process that should be adopted for the IRIS process. Finally, the committee offers some suggestions for improving the IRIS development process, providing a "roadmap" of the specific items for consideration.

An Overview of the Development of the Draft IRIS Assessment

In Chapter 2, the committee provided its own diagram (Figure 2-1) describing the steps used to generate the draft IRIS assessment. For the purpose of offering committee comments on ways to improve those steps, that figure has been expanded to indicate the key outcomes at each step (Figure 7-2). For each of the steps, the figure identifies the key questions addressed in the process. At the broadest level, the steps include systematic review of evidence, hazard identification using a weight-of-evidence approach, and dose-response assessment.

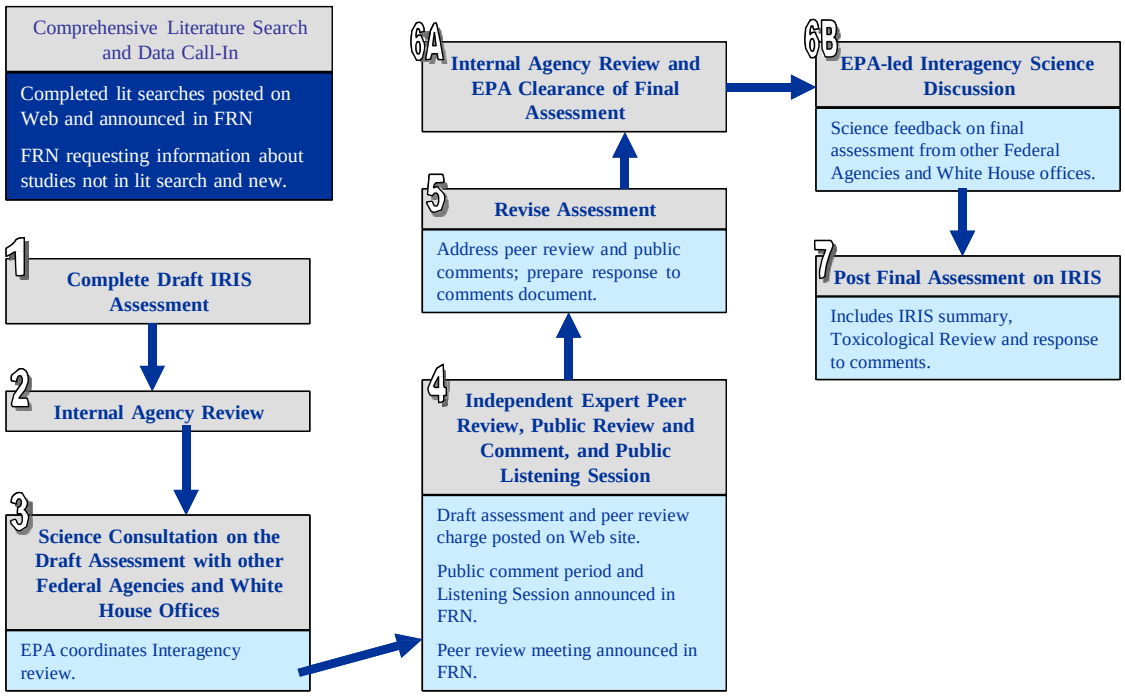

FIGURE 7-1 New IRIS assessment process. Abbreviations: FRN, Federal Reserve Note; IRIS, Integrated Risk Information System; EPA, Environmental Protection Agency. Source: EPA 2009a.

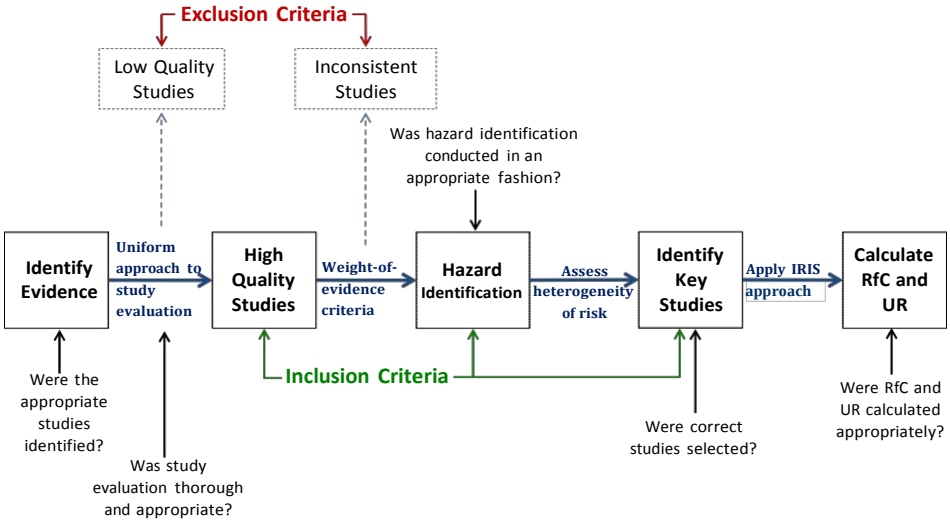

FIGURE 7-2 Elements of the key steps in the development of a draft IRIS assessment. Abbreviations: IRIS, Integrated Risk Information System; RfC, reference concentration; UR, unit risk.

The systematic review process is undertaken to identify all relevant literature on the agent of interest, to evaluate the identified studies, and possibly to provide a qualitative or quantitative synthesis of the literature. Chapter 1 of the draft IRIS assessment of formaldehyde provides a brief general description of the process followed by EPA, including the approach to searching the literature. However, neither Chapter 1 nor other chapters of the draft provide a sufficiently detailed description of the approach taken in evaluating individual studies. In discussing particular epidemiologic studies, a systematic approach to study evaluation is not provided. Consequently, some of the key methodologic points are inconsistently mentioned, such as information bias and confounding.

For hazard identification, the general guidance is also found in Chapter 1 of the draft IRIS assessment. The approach to conducting hazard identification is critical for the integrity of the IRIS process. The various guidelines cited in Chapter 1 provide a general indication of the approach to be taken to hazard identification but do not offer a clear template for carrying it out. For the formaldehyde assessment, hazard identification is particularly challenging because the outcomes include cancer and multiple noncancer outcomes. The various EPA guidelines themselves have not been harmonized, and they provide only general guidance. Ultimately, the quality of the studies reviewed and the strength of evidence provided by the studies for deriving RfCs and unit risks need to be clearly presented. More formulaic approaches are followed for calculation of RfCs and unit risks. The key issue is whether the calculations were conducted appropriately and according to accepted assessment procedures.

Brief Review of Established Best Practices

The following sections highlight some best practices of current approaches to evidence-based reviews, hazard identification, and dose-response assessment that could provide EPA guidance if it decides to address some of the fundamental issues identified by the committee. The discussion is meant not to be comprehensive or to provide all perspectives on the topics but simply to highlight some important aspects of the approaches. The committee recognizes that some of the concepts and approaches discussed below are elementary and are addressed in some of EPA's guidelines. However, the current state of the formaldehyde draft IRIS assessment suggests that there might be a problem with the practical implementation of the guidelines in completing the IRIS assessments. Therefore, the committee highlights aspects that it finds most critical.

Current Approaches to Evidence-Based Reviews

Public-health decision-making has a long history of using comprehensive reviews as the foundation for evaluating evidence and selecting policy options. The landmark 1964 report of the U.S. surgeon general on tobacco and disease is exemplary (DHEW 1964). It used a transparent method that involved a critical survey of all relevant literature by a neutral panel of experts and an explicit framework for assessing the strength of evidence for causation that was equivalent to hazard identification (Table 7-1).

The tradition of comprehensive, evidence-based reviews has been continued in the surgeon general's reports. The 2004 surgeon general's report, which marked the 40th anniversary of the first report, highlighted the approach for causal inference used in previous reports and provided an updated and standardized four-level system for describing strength of evidence (DHHS 2004) (Table 7-2).

The same systematic approaches have become fundamental in many fields of clinical medicine and public health. The paradigm of "evidence-based medicine" involves the systematic review of evidence as the basis of guidelines. The international Cochrane Collaboration engages thousands of researchers and clinicians throughout the world to carry out reviews. In the United States, the Agency for Healthcare Research and Quality supports 14 evidence-based practice centers to conduct reviews related to healthcare.

TABLE 7-1 Criteria for Determining Causality

Criterion	Definition
Consistency	Persistent association among different studies in different populations
Strength of association	Magnitude of the association
Specificity	Linkage of specific exposure to specific outcome
Temporality	Exposure comes before effect
Coherence, plausibility, analogy	Coherence of the various lines of evidence with a causal relationship
Biologic gradient	Presence of increasing effect with increasing exposure (dose-response relationship)
Experiment	Observations from "natural experiments," such as cessation of exposure (for example, quitting smoking)

Source: DHHS 2004.

TABLE 7-2 Hierarchy for Classifying Strength of Causal Inferences on the Basis of Available Evidence

A. Evidence is *sufficient* to infer a causal relationship.

B. Evidence is *suggestive but not sufficient* to infer a causal relationship.

C. Evidence is *inadequate* to infer the presence or absence of a causal relationship (evidence that is sparse, of poor quality, or conflicting).

D. Evidence is *suggestive of no causal relationship*.

Source: DHHS 2004.

There are also numerous reports from NRC committees and the Institute of Medicine (IOM) that exemplify the use of systematic reviews in evaluating evidence. Examples include reviews of the possible adverse responses associated with Agent Orange, vaccines, asbestos, arsenic in drinking water, and secondhand smoke. A 2008 IOM report, *Improving the Presumptive Disability Decision-Making Process for Veterans*, proposed a comprehensive new scheme for evaluating evidence that an exposure sustained in military service had contributed to disease (IOM 2008); the report offers relevant coverage of the practice of causal inference.

This brief and necessarily selective coverage of evidence reviews and evaluations shows that models are available that have proved successful in practice. They have several common elements: transparent and explicitly documented methods, consistent and critical evaluation of all relevant literature, application of a standardized approach for grading the strength of evidence, and clear and consistent summative language. Finally, highlighting features and limitations of the studies for use in quantitative assessments seems especially important for IRIS literature reviews.

A state-of-the-art literature review is essential for ensuring that the process of gathering evidence is comprehensive, transparent, and balanced. The committee suggests that EPA develop a detailed search strategy with search terms related to the specific questions that are addressed by the literature review. The yield of articles from searches can best be displayed graphically, documenting how initial search findings are narrowed to the articles in the final review selection on the basis of inclusion and exclusion criteria. Figure 7-3 provides an example of the selection process in a systematic review of a drug for lung disease. The progression from the initial 3,153 identified articles to the 11 reviewed is transparent. Although this example comes from an epidemiologic meta-analysis, a similar transparent process in which search terms, databases, and resources are listed and study selection is carefully tracked may be useful at all stages of the development of the IRIS assessment.

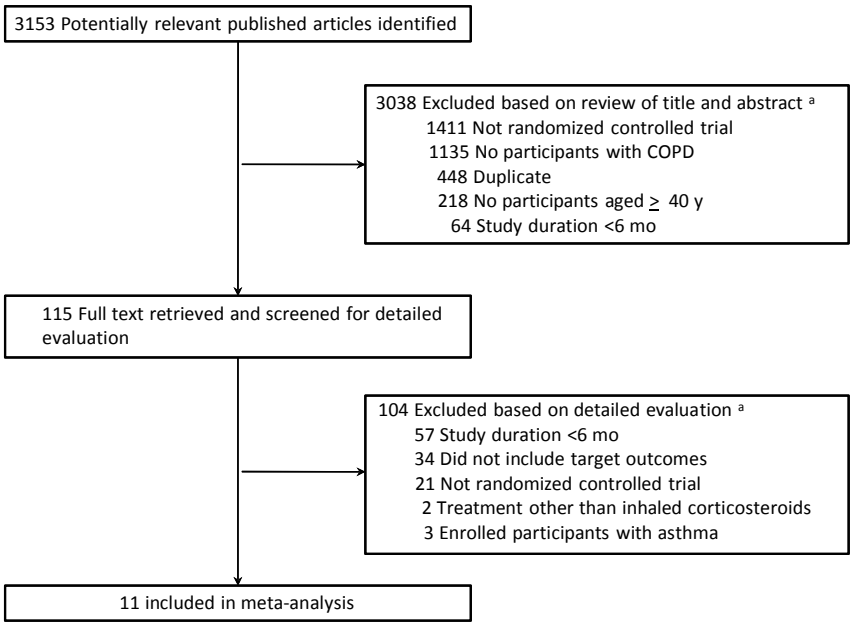

FIGURE 7-3 Example of an article-selection process. [a]Articles could be excluded for more than one reason; therefore, summed exclusions exceed total. Abbreviations: COPD, chronic obstructive pulmonary disease. Source: Drummond et al. 2008. Reprinted with permission; copyright 2008, American Medical Association.

After studies are identified for review, the next step is to summarize the details and findings in evidence tables. Typically, such tables provide a link to the references, details of the study populations and methods, and key findings. They are prepared in a rigorous fashion with quality-assurance measures, such as using multiple abstractors (at least for a sample) and checking all numbers abstracted. If prepared correctly, the tables eliminate the need for long descriptions of studies and result in shorter text. Some draft IRIS assessments have begun to use a tabular format for systematic and concise presentation of evidence, and the committee encourages EPA to refine and expand that format as it revises the formaldehyde draft IRIS assessment and begins work on others.

The methods and findings of the studies are then evaluated with a standardized approach. Templates are useful for this purpose to ensure uniformity of approach, particularly if multiple reviewers are involved. Such standardized approaches are applied whether the research is epidemiologic (observational), experimental (randomized clinical trials), or toxicologic (animal bioassays). For example, for an observational epidemiologic study, a template for evaluation should consider the following:

- Approach used to identify the study population and the potential for selection bias.
- Study population characteristics and the generalizability of findings to other populations.
- Approach used for exposure assessment and the potential for information bias, whether differential (nonrandom) or nondifferential (random).
- Approach used for outcome identification and any potential bias.
- Appropriateness of analytic methods used.
- Potential for confounding to have influenced the findings.
- Precision of estimates of effect.
- Availability of an exposure metric that is used to model the severity of adverse response associated with a gradient of exposures.

Similarly, a template for evaluation of a toxicology study in laboratory animals should consider the species and sex of animals studied, dosing information (dose spacing, dose duration, and route of exposure), end points considered, and the relevance of the end points to human end points of concern.

Current Approaches to Hazard Identification

Hazard identification involves answering the question, Does the agent cause the adverse effect? (NRC 1983, 2009). Numerous approaches have been used for this purpose, and there is an extensive literature on causal inference, both on its philosophic underpinnings and on methods for evaluating the strength of evidence of causation. All approaches have in common a systematic identification of relevant evidence, criteria for evaluating the strength of evidence, and language for describing the strength of evidence of causation. The topic of causal inference and its role in decision-making was recently covered in the 2008 IOM report on evaluation of the presumptive decision-making process noted above. The 2004 report of the U.S. surgeon general on smoking and health (DHHS 2004) provided an updated review of the methods used in that series of reports.

The review approach for hazard identification embodies the elements described above and uses the criteria for evidence evaluation that have their origins in the 1964 report of the U.S. surgeon general (DHEW 1964) and the writings of Austin Bradford Hill, commonly known as the Hill criteria (see Table 7-1; Hill 1965). The criteria are not rigid and are not applied in a check-list manner; in fact, none is required for inferring a causal relationship, except for temporality inasmuch as exposure to the causal agent must precede the associated effect. The conclusion of causal inference is a clear statement on the strength of evidence of causation. For the purpose of hazard identification, such statements should follow a standardized classification to avoid ambiguity and to ensure comparability among different agents and outcomes.

Beyond the surgeon general's reports used here as an example, there are numerous examples of systematic approaches to hazard identification, including the monographs on carcinogenicity of the International Agency for Research on Cancer and the National Toxicology Program.[1] They have the same elements of systematic gathering and review of all lines of evidence and classification of the strength of evidence in a uniform and hierarchic structure.

Current Approaches to Dose-Response Assessment

The topic of dose-response assessment was covered in *Science and Decisions* (NRC 2009), which reviewed the current paradigm and called for a unified framework, bringing commonality to approaches for cancer and noncancer end points. That report also provides guidance on enhancing methods used to characterize uncertainty and variability. The present committee supports those recommendations but offers additional suggestions on the complementary coverage of the use of meta-analysis and pooled analysis in dose-response assessment.

IRIS assessments should address the following critical questions: Which studies should be included for derivation of reference values for noncancer outcomes and unit risks for cancer outcomes? Which dose-response models should be used for deriving those values? The latter question is related to model uncertainty in quantitative risk assessment and is not addressed here in this report. The former question is related to a fundamental issue of filtering the literature to identify the studies that provide the best dose-response information. A related question arises about how to combine information among studies because multiple studies may provide sufficient dose-response data. For this section, the committee assumes that the previously described evidence-based review has identified studies with adequate dose-response information to support some quantification of risk associated with exposure.

As suggested above, it would be unusual for a single study to trump all other studies providing information for setting reference values and unit risks. The combination of the analysis outcomes of different studies falls under the general description of meta-analysis (Normand 1999). The combination and synthesis of results of different studies appears central to an IRIS assessment, but such analyses require careful framing.

Stroup and colleagues (2000) provide a summary of recommendations for reporting meta-analyses of epidemiologic studies. Their proposal includes a table with a proposed check list that has broad categories for reporting, including background (such as problem definition and study population), search strategy (such as searchers, databases, and registries used), methods, results (such as graphic and tabular summaries, study description, and statistical uncertainty), discussion (such as bias and quality of included studies), and conclusion (such as generalization of conclusions and alternative explanations). Their recommendations on methods warrant specific consideration with reference to the development of an IRIS assessment, particularly those on evaluation and assessment of study relevance, rationale for selection and coding of studies, confounding, study quality, heterogeneity, and statistical methods. For the latter, key issues include the selection of models, the clarity with which findings are presented, and the availability of sufficient details to facilitate replication.

In combining study information, it is important that studies provide information on the same quantitative outcome, are conducted under similar conditions, and are of similar quality. If studies are of different quality, this might be addressed by weighting.

The simplest form of combining study information involves the aggregation of p values among a set of independent studies of the same null hypothesis. That simple approach might have appeal for establishing the relationship between some risk factor and an adverse outcome, but it is not useful for establishing exposure levels for a hazard. Thus, effect-size estimation among studies is usually of more interest for risk-estimation purposes and causality assessment. In this situation, a given effect is estimated for each study, and a combined estimate is obtained as a weighted average of study-specific effects in

[1]See http://monographs.iarc.fr/index.php and http://ntp.niehs.nih.gov/.

which the weights are inversely related to the precision associated with the estimation of each study-specific effect.

The question is whether EPA should routinely conduct meta-analysis for its IRIS assessments. Implicitly, the development of an IRIS assessment involves many of the steps associated with meta-analysis, including the collection and assessment of background literature. Assuming the availability of independent studies of the same end point and a comprehensive and unbiased inclusion of studies, questions addressed by a meta-analysis may be of great interest. Is there evidence of a homogeneous effect among studies? If not, can one understand the source of heterogeneity? If it is determined that a combined estimate is of interest (for example, an estimate of lifetime cancer risk based on combining study-specific estimates of this risk), a weighted estimate might be derived and reported.

Case Study: Revision of the Approach to Evidence Review and Risk Assessment for National Ambient Air Quality Standards

Approaches to evidence review and risk assessment vary within EPA. The recently revised approach used for NAAQSs offers an example that is particularly relevant because it represents a major change in an approach taken by one group in the National Center for Environmental Assessment. (EPA 2009b, 2010a,b)

Under Section 109 of the Clean Air Act, EPA is required to consider revisions of the NAAQSs for specified criteria air pollutants—currently particulate matter (PM), ozone, nitrogen dioxide, sulfur dioxide, carbon monoxide, and lead—every 5 years. Through 2009, the process for revision involved the development of two related documents that were both reviewed by the Clean Air Scientific Advisory Committee (CASAC) and made available for public comment. The first, the criteria document, was an encyclopedic compilation, sometimes several thousand pages long, of most scientific publications on the criteria pollutant that had been published since the previous review. Multiple authors contributed to the document, and there was generally little synthesis of the evidence, which was not accomplished in a systematic manner.

The other document was referred to as the staff paper. It was written by a different team in the Office of Air Quality Policy and Standards, and it identified the key scientific advances in the criteria document that were relevant to revising the NAAQSs. In the context of those advances, it offered the array of policy options around retaining or revising the NAAQSs that could be justified by recent research evidence. The linkages between the criteria document and the staff paper were general and not transparent.

The identified limitations of the process led to a proposal for its revision, and it took 2 years to complete the changes in the process. The new process replaces the criteria document with an integrated science assessment and a staff paper that includes a policy assessment. For the one pollutant, PM, that has nearly completed the full sequence, a risk and exposure analysis was also included.

The new documents address limitations of those used previously. The integrated science assessment is an evidence-based review that targets new studies as before. However, review methods are explicitly stated, and studies are reviewed in an informative and purposeful manner rather than in encyclopedic fashion. A main purpose of the integrated science assessment is to assess whether adverse health effects are causally linked to the pollutant under review. The integrated science assessment offers a five-category grading of strength of evidence on each outcome and follows the general weight-of-evidence approaches long used in public health. The intent is to base the risk and exposure analysis on effects for which causality is inferred or those at lower levels if they have particular public-health significance. The risk and exposure analysis brings together the quantitative information on risk and exposure and provides estimates of the current burden of attributable morbidity and mortality and the estimates of avoidable and residual morbidity and mortality under various scenarios of changes in the NAAQS. Standard descriptors for uncertainty are now in place.

The policy assessment develops policy options on the basis of the findings of the integrated science assessment and the risk and exposure analysis. The policy assessment for the PM NAAQS is framed around a series of policy-relevant questions, such as, Does the available scientific evidence, as reflected in the integrated science assessment, support or call into question the adequacy of the protection afforded by the current 24-hr PM_{10} standard against effects associated with exposures to thoracic coarse particles? Evidence-based answers to the questions are provided with a reasonably standardized terminology for uncertainty.

For the most recent reassessment of the PM NAAQS, EPA staff and CASAC found the process to be effective; it led to greater transparency in evidence review and development of policy options than the prior process (Samet 2010). As noted above, the present committee sees the revision of the NAAQS review process as a useful example of how the agency was able to revise an entrenched process in a relatively short time.

Reframing the Development of the IRIS Assessment

The committee was given the broad charge of reviewing the formaldehyde draft IRIS assessment and also asked to consider some specific questions. In addressing those questions, the committee found, as documented in Chapter 2, that some problems with the draft arose because of the processes and methods used to develop the assessment. Other committees have noted some of the same problems. Accordingly, the committee suggests here steps that EPA could take to improve IRIS assessment through the implementation of methods that would better reflect current practices. The committee offers a roadmap for changes in the development process if EPA concludes that such changes are needed. The term *roadmap* is used because the topics that need to be addressed are set out, but detailed guidance is not provided because that is seen as beyond the committee's charge. The committee's discussion of a reframing of the IRIS development process is based on its generic representation provided in Figure 7-2. The committee recognizes that the changes suggested would involve a multiyear process and extensive effort by the staff of the National Center for Environmental Assessment and input and review by the EPA Science Advisory Board and others. The recent revision of the NAAQS review process provides an example of an overhauling of an EPA evidence-review and risk-assessment process that took about 2 years.

In the judgment of the present and past committees, consideration needs to be given to how each step of the process could be improved and gains made in transparency and efficiency. Models for conducting IRIS reviews more effectively and efficiently are available. For each of the various components (Figure 7-2), methods have been developed, and there are exemplary approaches in assessments carried out elsewhere in EPA and by other organizations. In addition, there are relevant examples of evidence-based algorithms that EPA could draw on. Guidelines and protocols for the conduct of evidence-based reviews are available, as are guidelines for inference as to the strength of evidence of association and causation. Thus, EPA may be able to make changes in the assessment process relatively quickly by drawing on appropriate experts and selecting and adapting existing approaches.

One major, overarching issue is the use of weight of evidence in hazard identification. The committee recognizes that the terminology is embedded in various EPA guidelines (see Appendix B) and has proved useful. The determination of weight of evidence relies heavily on expert judgment. As called for by others, EPA might direct effort at better understanding how weight-of-evidence determinations are made with a goal of improving the process (White et al. 2009).

The committee highlights below what it considers critical for the development of a scientifically sound IRIS assessment. Although many elements are basic and have been addressed in the numerous EPA guidelines, implementation does not appear to be systematic or uniform in the development of the IRIS assessments.

General Guidance for the Overall Process

- Elaborate an overall, documented, and quality-controlled process for IRIS assessments.
- Ensure standardization of review and evaluation approaches among contributors and teams of contributors; for example, include standard approaches for reviews of various types of studies to ensure uniformity.
- Assess disciplinary structure of teams needed to conduct the assessments.

Evidence Identification: Literature Collection and Collation Phase

- Select outcomes on the basis of available evidence and understanding of mode of action.
- Establish standard protocols for evidence identification.
- Develop a template for description of the search approach.
- Use a database, such as the Health and Environmental Research Online (HERO) database, to capture study information and relevant quantitative data.

Evidence Evaluation: Hazard Identification and Dose-Response Modeling

- Standardize the presentation of reviewed studies in tabular or graphic form to capture the key dimensions of study characteristics, weight of evidence, and utility as a basis for deriving reference values and unit risks.
- Develop templates for evidence tables, forest plots, or other displays.
- Establish protocols for review of major types of studies, such as epidemiologic and bioassay.

Weight-of-Evidence Evaluation: Synthesis of Evidence for Hazard Identification

- Review use of existing weight-of-evidence guidelines.
- Standardize approach to using weight-of-evidence guidelines.
- Conduct agency workshops on approaches to implementing weight-of-evidence guidelines.
- Develop uniform language to describe strength of evidence on noncancer effects.
- Expand and harmonize the approach for characterizing uncertainty and variability.
- To the extent possible, unify consideration of outcomes around common modes of action rather than considering multiple outcomes separately.

Selection of Studies for Derivation of Reference Values and Unit Risks

- Establish clear guidelines for study selection.
  - Balance strengths and weaknesses.
  - Weigh human vs experimental evidence.
  - Determine whether combining estimates among studies is warranted.

Calculation of Reference Values and Unit Risks

- Describe and justify assumptions and models used. This step includes review of dosimetry models and the implications of the models for uncertainty factors; determination of appropriate points of

departure (such as benchmark dose, no-observed-adverse-effect level, and lowest observed-adverse-effect level), and assessment of the analyses that underlie the points of departure.

- Provide explanation of the risk-estimation modeling processes (for example, a statistical or biologic model fit to the data) that are used to develop a unit risk estimate.
- Assess the sensitivity of derived estimates to model assumptions and end points selected. This step should include appropriate tabular and graphic displays to illustrate the range of the estimates and the effect of uncertainty factors on the estimates.
- Provide adequate documentation for conclusions and estimation of reference values and unit risks. As noted by the committee throughout the present report, sufficient support for conclusions in the formaldehyde draft IRIS assessment is often lacking. Given that the development of specific IRIS assessments and their conclusions are of interest to many stakeholders, it is important that they provide sufficient references and supporting documentation for their conclusions. Detailed appendixes, which might be made available only electronically, should be provided when appropriate.

REFERENCES

DHEW (U.S. Deaprtment of Health Education and Welfare). 1964. Smoking and Health. Report of the Advisory Committee to the Surgeon General. Public Health Service Publication No. 1103. Washington, DC: U.S. Government Printing Office [online]. Available: http://profiles.nlm.nih.gov/NN/B/B/M/Q/_/nnbbmq.pdf [accessed Feb. 1, 2011].

DHHS (U.S. Department of Health and Human Services). 2004. The Health Consequences of Smoking: A Report of the Surgeon General. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, Atlanta, GA [online]. Available: http://www.cdc.gov/tobacco/data_statistics/sgr/2004/complete_report/index.htm [accessed Nov. 22, 2010].

Drummond, M.B., E.C. Dasenbrook, M.W. Pitz, D.J. Murphy, and E. Fan. 2008. Inhaled corticosteroids in patients with stable chronic obstructive pulmonary disease: A systematic review and meta-analysis. JAMA. 300(20):2407-2416.

EPA (U.S. Environmental Protection Agency). 2009a. New Process for Development of Integrated Risk Information System Health Assessments. Memorandum to Assistant Administrators, General Counsel, Inspector General, Chief Financial Officer, Chief of Staff, Associate Administrators, and Regional Administrators, from Lisa P. Jackson, the Administrator, U.S. Environmental Protection Agency, Washington, DC. May 21, 2009 [online]. Available: http://www.epa.gov/iris/pdfs/IRIS_PROCESS_MEMO.5.21.09.PDF [accessed Nov. 23, 2010].

EPA (U.S. Environmental Protection Agency). 2009b. Integrated Science Assessment for Particulate Matter (Final Report). EPA/600/R-08/139F. National Center for Environmental Assessment-RTP Division, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC. December 2009 [online]. Available: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=216546 [accessed March 2, 2011].

EPA (U.S. Environmental Protection Agency). 2010a. Quantitative Health Risk Assessment for Particulate Matter (Final Report). EPA-452/R-10-005. Office of Air Quality Planning and Standards, Office of Air and Radiation, U.S. Environmental Protection Agency, Research Triangle Park, NC. June 2010 [online]. Available: http://www.epa.gov/ttn/naaqs/standards/pm/data/PM_RA_FINAL_June_2010.pdf [accessed March 2, 2011].

EPA (U.S. Environmental Protection Agency). 2010b. Policy Assessment for the Review of the Particulate Matter National Ambient Air Quality Standards (Second External Review Draft). EPA 452/P-10-007. Office of Air Quality Planning and Standards, Office of Air and Radiation, U.S. Environmental Protection Agency, Research Triangle Park, NC. June 2010 [online]. Available: http://www.epa.gov/ttnnaaqs/standards/pm/data/20100630seconddraftpmpa.pdf [accessed March 2, 2011].

Hill, A.B. 1965. The environment and disease: Association or causation? Proc. R. Soc. Med. 58:295-300

IOM (Institute of Medicine). 2008. Improving the Presumptive Disability Decision-Making Process for Veterans. Washington, DC: National Academies Press.

NRC (National Research Council). 1983. Risk Assessment in the Federal Government: Managing the Process. Washington, DC: National Academy Press.

NRC (National Research Council). 2009. Science and Decisions: Advancing Risk Assessment. Washington, DC: National Academies Press.

Normand, S.L. 1999. Meta-analysis: Formulating, evaluating, combining, and reporting. Stat. Med. 18(3): 321-359.

Samet, J.M. 2010. CASAC Review of Policy Assessment for the Review of the PM NAAQS - Second External Review Draft (June 2010). EPA-CASAC-10-015. Letter to Lisa P. Jackson, Administrator, from Jonathan M. Samet, Clean Air Scientific Advisory Committee, Office of Administrator, Science Advisory Board, U.S. Environmental Protection Agency, Washington, DC. September 10, 2010 [online]. Available: http://yosemite.epa.gov/sab/sabproduct.nsf/CCF9F4C0500C500F8525779D0073C593/$File/EPA-CASAC-10-015-unsigned.pdf [accessed Nov. 22, 2010].

Stroup, D.F., J.A. Berlin, S.C. Morton, I. Olkin, G.D. Williamson, D. Rennie, D. Moher, B.J. Becker, T.A. Sipe, and S.B. Thacker. 2000. Meta-analysis of observational studies in epidemiology: A proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA. 283(15):2008-2012.

White, R.H., I. Cote, L. Zeise, M. Fox, F. Dominici, T.A. Burke, P.D. White, D.B. Hattis, and J.M. Samet. 2009. State-of-the-science workshop report: Issues and approaches in low-dose-response extrapolation for environmental health risk assessment. Environ. Health Perspect. 117(2):283-287.